Supporting Information

Simultaneous NMR study of protein structure and dynamics using conservative mutagenesis

Laboratory of Chemical Physics, NIDDK, National Institutes of Health, Bethesda, MD 20892-0520

RECEIVED DATE (automatically inserted by publisher)

Lishan Yao,[a] Beat Voegeli,[a] Dennis A. Torchia,[b] and Ad Bax[a]
[a] Laboratory of Chemical Physics, NIDDK
[b] National Institute of Dental and Cranofacial Research
National Institutes of Health, Bethesda, MD 20892-0520

Table S1. The five elements of alignment vector *a* and the corresponding D_a, *R*, and Euler angles (α, β, γ) used in the RDC simulations.

	K19AD47K	K19ED40N	K19EK4A-C-His_6	K19EK4A-N-His_6	K19AT11K
a_1	-13.19	7.235	9.935	0.822	-6.513
a_2	-15.65	-7.600	0.602	-11.736	-8.903
a_3	-3.576	-1.334	-1.716	2.097	-2.252
a_4	5.247	-3.996	3.146	11.867	6.921
a_5	-1.000	10.11	4.296	7.231	11.487
D_a	10.44	-7.292	5.226	8.933	8.376
R	0.287	0.330	0.508	0.264	0.352
α	272.7	153.2	111.1	75.4	65.8
β	98.1	88.6	9.54	58.6	78.7
γ	249.9	44.7	108.9	8.0	141.0

Table S2. The dynamics parameters used in generating the synthetic RDCs.

Residue	*S*	η	γ(°)
3	0.976	0.013	88.7
4	0.951	0.028	148.0
5	0.978	0.01	13.9
6	0.967	0.016	96.8
7	0.942	0.03	57.3
8	0.951	0.025	147.4
9	0.932	0.048	167.5
10	0.911	0.031	114.5
11	0.807	0.072	157.4
12	0.705	0.131	33.6
13	0.866	0.038	133.4
14	0.824	0.076	139.6
15	0.859	0.037	154.3
16	0.934	0.036	30.7
17	0.894	0.053	61.7
18	0.937	0.017	160.5
19	0.931	0.036	168.3
20	0.881	0.041	47.9
21	0.928	0.035	2.7
22	0.957	0.029	11.4
23	0.959	0.01	61.1
24	0.956	0.022	66.5
25	0.93	0.046	111.6
26	0.977	0.01	80.7
27	0.956	0.019	99.9
28	0.938	0.019	31.6
29	0.98	0.011	97.8
30	0.973	0.017	131.1

31	0.977	0.011	53.8
32	0.954	0.031	149.1
33	0.973	0.01	8.2
34	0.98	0.006	81.8
35	0.953	0.021	1.1
36	0.958	0.013	133.4
37	0.928	0.043	125.7
38	0.933	0.03	38.1
39	0.953	0.02	30.3
40	0.801	0.12	58.6
41	0.646	0.17	23.2
42	0.928	0.026	116.5
43	0.957	0.018	45.3
44	0.954	0.024	28.0
45	0.912	0.055	110.6
46	0.957	0.022	130.1
47	0.85	0.069	155.1
48	0.957	0.018	1.1
49	0.983	0.006	139.1
50	0.938	0.023	5.4
51	0.976	0.014	55.8
52	0.949	0.022	82.2
53	0.942	0.019	20.7
54	0.868	0.038	136.6
55	0.955	0.029	11.1
56	0.954	0.024	119.1

Table S3. The dynamics parameters of GB3 N-H vectors extracted from 6 alignments.

Residue	*S*	*S* error	η	η error	γ	γ error
3	0.959	0.016				
4	0.947	0.019				
5	0.937	0.033				
6	0.949	0.028				
7	0.948	0.033				
8	0.927	0.03				
9	0.917	0.036				
10	0.770	0.033	0.113	0.072	5.39	52.197
12	0.522	0.058	0.554	0.145	4.495	2.721
13	0.866	0.03				
14	0.821	0.047	0.079	0.043	8.244	11.71
15	0.842	0.049	0.083	0.037	6.662	13.244
16	0.899	0.037	0.051	0.026	1.702	54.083
17	0.882	0.023	0.05	0.025	8.339	21.903
18	0.920	0.033	0.067	0.034	25.888	34.496
20	0.827	0.061	0.187	0.089	64.705	13.035
21	0.857	0.056	0.107	0.055	51.951	20.544
22	0.961	0.022	0.068	0.033	19.263	26.165
23	0.922	0.019				
24	0.907	0.032				
26	0.970	0.015				
27	0.959	0.031				
28	0.929	0.023				
29	0.957	0.017				
31	0.960	0.044				
32	0.903	0.019				
33	0.940	0.019				
34	0.925	0.039				
35	0.936	0.029				
36	0.934	0.02				
37	0.867	0.027	0.091	0.047	-33.02	33.452
38	1.019	0.04	0.048	0.029	39.837	52.106
39	0.904	0.022				
42	0.964	0.038	0.041	0.022	-36.717	72.176
43	0.942	0.03				
44	1.041	0.063	0.067	0.028	22.982	28.626
45	0.900	0.042	0.049	0.02	8.68	22.052
46	0.941	0.032				
47	0.828	0.056	0.091	0.055	81.31	27.156
48	0.808	0.039	0.138	0.102	19.032	30.706
49	0.893	0.026	0.044	0.044	15.913	26.595
50	0.906	0.036	0.033	0.019	7.699	72.482
51	0.928	0.027	0.054	0.028	21.376	21.443
52	0.919	0.039				
53	0.976	0.03				
54	0.983	0.032				
55	0.935	0.031				
56	0.937	0.037				

Table S4. The dynamics parameters of GB3 C^{α}-H^{α} vectors extracted from 6 alignments.

Residue	*S*	*S* error	η	η error
2	0.952	0.022		
3	0.941	0.02		
4	0.914	0.025		
5	0.947	0.048	0.024	0.021
6	0.923	0.033		
7	0.919	0.026		
8	0.810	0.065	0.095	0.05
10	0.836	0.018		
12	0.785	0.066	0.098	0.055
13	0.732	0.066	0.115	0.05
15	0.897	0.05	0.12	0.052
16	0.963	0.046	0.054	0.031
17	0.893	0.032		
18	0.974	0.029	0.048	0.025
20	0.808	0.097	0.12	0.124
21	0.898	0.026		
22	0.921	0.042	0.127	0.065
23	0.914	0.028		
24	0.931	0.019		
26	0.914	0.026		
27	0.904	0.026	0.069	0.033
28	0.935	0.017		
29	0.954	0.022		
31	0.973	0.026	0.052	0.032
32	0.954	0.026		
33	0.917	0.017		
34	0.961	0.031		
35	0.937	0.031		
36	0.889	0.031		
37	0.901	0.02	0.075	0.033
39	0.924	0.017		
42	0.989	0.027	0.086	0.03
43	0.940	0.03		
44	0.905	0.025		
45	0.955	0.028		
46	0.906	0.039	0.058	0.037
47	1.000	0.02		
48	0.861	0.026	0.168	0.068
49	0.972	0.022		
50	0.905	0.032		
51	0.906	0.034	0.101	0.046
52	0.938	0.031		
53	0.931	0.032		
54	0.924	0.028		
55	0.935	0.032		

Table S5. The dynamics parameters of GB3 C^{α}-C′ vectors extracted from 6 alignments.

Residue	*S*	*S* error
2	0.967	0.087
3	0.903	0.203
4	0.985	0.105
5	0.999	0.142
6	0.981	0.067
7	1.049	0.055
8	1.006	0.082
9	0.861	0.131
10	1.017	0.094
12	0.866	0.055
13	0.908	0.072
14	0.879	0.144
15	0.973	0.163
16	0.917	0.181
17	0.975	0.073
18	0.900	0.07
20	0.934	0.073
21	0.884	0.156
22	0.968	0.074
23	1.016	0.075
24	0.950	0.043
26	0.979	0.036
27	0.964	0.106
28	0.940	0.096
29	0.907	0.107
31	0.997	0.083
32	0.921	0.105
33	0.977	0.037
34	0.967	0.08
35	1.025	0.12
36	0.804	0.086
37	0.957	0.079
38	0.955	0.059
39	0.827	0.072
42	1.001	0.033
43	0.974	0.054
44	0.964	0.06
45	0.978	0.074
46	0.985	0.045
47	0.982	0.071
48	1.028	0.075
49	0.981	0.058
50	0.901	0.09
51	0.992	0.103
52	0.992	0.111
53	0.992	0.111
54	1.121	0.093
55	1.058	0.102

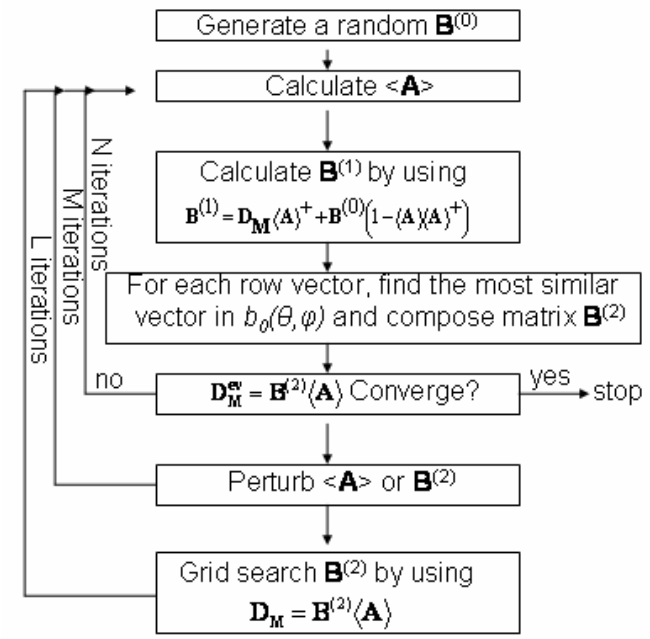

Figure S1. Flowchart of the iterative DIDC method. The two dimensional $b_0(\theta_i,\phi_j)$ grid was generated with 0.2° grid size for both θ and ϕ. When θ is close to 0°, the grid becomes finer, requiring more computer time per unit of solid angle. To solve this problem, the grid size was reduced as follows: The grid was generated starting from θ = 0° and ϕ = 0°, ϕ was increased by 0.2° to 360° and then θ was increased by 0.2° to 90°. For each generated grid point (except the first one), the following procedure was used to accept or reject this point by examining the distance between two b_0 grid vectors defined as $\left|b_{0i}(\theta,\phi)-b_{0j}(\theta,\phi)\right|$, where $b_{0j}(\theta,\phi)$ is the current grid point and $b_{0i}(\theta,\phi)$ is one prior accepted grid point. If the distance between the current point and any accepted grid point is smaller than 0.007, $b_{0j}(\theta,\phi)$ is excluded; otherwise, $b_{0j}(\theta,\phi)$ is accepted. The error caused by the grid is rather small. For example the error for the five sets of simulated N-H RDCs caused by the grid is about 0.03 Hz, 5-10 times smaller than typical experimental error. In the first loop of the flowchart (box 5, N iterations), a local minimum can be found in approximately ~10-20 iterations and for computational efficiency N is set to 10. If the fitting error is smaller than the predefined tolerance the algorithm stops; otherwise it goes to box 6, the Monte Carlo search. During the Monte Carlo search, the <**A**> and **B** are perturbed alternately for odd- and even-numbered M values. First, the perturbation of <**A**> is carried out by adding a small random number (between -1 and 1) to each elements of the <**A**> matrix. Second, the algorithm goes to box 2 and the first iteration (from box 5 to 2) is carried out to find a new local minimum. This local minimum is compared to the previous local minimum by using the Metropolis criterion $e^{(Ep-Ec)/T}$ where *Ec*, *Ep* is the current and previous RDCs fitting error (total χ^2, summed over all RDCs in all alignments) and T is the temperature factor (T is set to 4). If $e^{(Ep-Ec)/T}$ is smaller than a random number (0< random number <1) or *Ec* < *Ep*, this perturbation is accepted, otherwise rejected. Third, the perturbation of **B** is carried out. A random angle (from -20° to 20°) is added to ϕ of vectors with the three largest RDC

fitting errors. The algorithm goes to box 2 searching for a new minimum and the same procedure is used to accept or reject the move as in the <**A**> perturbation. M is set to 20. It turns out that the perturbation of <**A**> generally has higher acceptance ratio and appears more effective. The <**A**> with the smallest fitting RDC error is selected and the algorithm goes to the next box (box 7). L is set to 6. When L>3, the temperature is decreased to 0.7 and the perturbation of <**A**> is reduced (to a random number between -0.2 and 0.2) while the perturbation of **B** is not carried out. The algorithm is written in Matlab. It takes ~3 minutes to run the program in our test on an 80 (vectors) × 6 (alignments) RDC matrix using a Pentium 3.4G Hz dual core Windows machine. For the sets of RDCs simulated for three alignments, N = 10; M = 40; L = 20, and the random structures are used as the initial guesses.

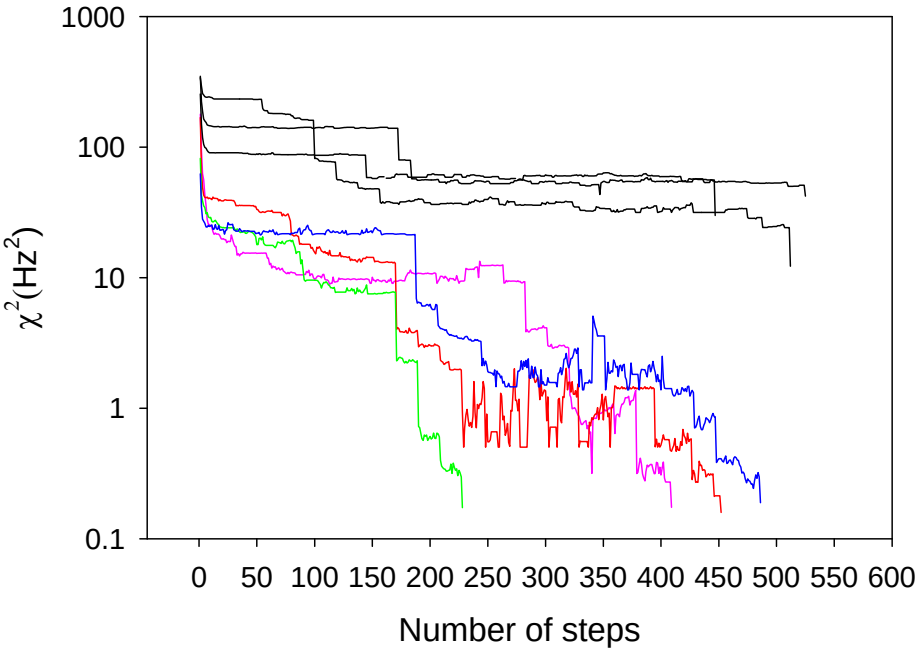

Figure S2. χ^2 convergence trajectories of the iterative DIDC method for four converging (colored lines) and three non-converging (black lines) structures, using RDCs from three alignments and no added noise, starting from random vector orientations. The vertical axis represents the total fitting error χ^2 ($\Sigma(RDC^{syn}-RDC^{calc})^2$) (prior to box 6 in Scheme 1, main text) and the horizontal axis represents the number of Monte Carlo steps used in the iterative DIDC method. Starting values for χ^2, i.e., the first time when entering box 5 in Scheme 1, are *ca* 10^4 Hz^2. For the final step, two values of χ^2 are shown, the value prior to box 6 and the value after the grid search of box 7.